INTRODUCTION

Go to your bookshelf and pull off a book; any book. It may be one you have read many times, or it could be one that is still on your "to read" list. Take a look at it. It may be bound with the flimsy cardboard of a paperback sporting a slick, shiny cover. Or the pages could be held between the cloth-covered boards of a quality hardback. It is unlikely, however, unless you are either very wealthy or very lucky, that your fingers will be touching a fine leather binding.

It is probable that you did not purchase the book for its physical appearance, as appealing as that may be, but for its content. That's where things get complicated in our story: complicated because it is very hard to say what the content consists of. Words, yes, but you didn't buy just a set of words, unless your book is a dictionary. No, you bought this book for the story it tells or for the information it imparts. You may have been seeking entertainment, or to learn something new (and happy is the person who gets both!). Although the story or the information came to you as words, you may not be able to recite even a small passage verbatim. We read the words but we remember the meaning, another concept that is difficult to define.

If I ask you some questions about the book, some will be easy to answer, some more difficult. I could ask you for the title, and most likely you know that. The same for the author. You could surely tell me what the book is about, either with a topic ("it's a history of the Venetian Republic") or a story ("it's about a girl who lives on the prairie and what she and her family go through to survive"). Chances are, though, that if I ask you who published the book, you'll be taking a sneak peek at the title page or the spine to find that information. The place and date of publication will not only be less imprinted in your consciousness but they may actually be a bit hard to find. The precise number of pages is another undeniable fact about the book that may not be on the tip of your tongue.

As a reader, it is the reading experience and what it leaves behind in your memory that makes up the inherent value of the book. And we do know that

readers value their books. There would be no other reason to use the bulk of the wall space of one's home for shelves for book storage, or, when moving to a new home, to pack, lug, and unpack untold pounds of what appears to be inert tree pulp.

Now let's leave books behind and look at other media. Just as many of us love our books, we also have among us many music lovers; people with towering racks of CDs or digital devices chock full of tunes. Here, though, we find some differences from our book story. Ask a music lover the "author" (composer) of a tune and you may be asking the obvious ("Beethoven's 5th symphony") or not ("Santa Claus Is Coming to Town"), even though both pieces of music are easily recognizable when heard. They are recognizable also because, unlike books, we listen to the same piece of music many times, and in different versions. This is a function of the fact that music is performed. Some performances are faithful interpretations of the music, and others, like jazz or digital sampling, are creative distortions of the original.

Music lovers with sufficient talent can reproduce a version of the music either by humming, singing, or playing the music on an instrument. We remember the notes of music in a way that we do not remember the words of a book. But if asked what the music is "about" we are in some difficulty in most cases. Unless the music has a specific story attached to it, such as Sergei Prokofiev's *Peter and the Wolf,* or the teen drama of "Dead Man's Curve," much music does not have a plot or a message that translates to "aboutness."

Other information that only dedicated aficionados of a music genre can relate about their listening choices are date of recording; names of all performers; date of composition; number and types of instruments. Asked what type of music we like, the answers are broad categories like rock, jazz, classical, or country; or sometimes a more specific category, still covering a wide swath: heavy metal; mostly Mozart; Irish folk music; Reggae.

Books and music are two common creative forms that many of us encounter in our everyday lives, and yet what we know about them and how we interact with them are quite different. Now let's look at another creative form: computer games. A player will know the name of the game, the general plot of the play (capture castle, defeat enemy, solve puzzle), and the names of characters. She will also know what capabilities she has as a player (running, jumping, opening doors). If it is a multi-player game, she will know the names of other players—that is, the names they are using in the game. She may not, however, be able to respond to the question "who wrote or created the game?" Games often do not have measurable durations although some have ending points, so asking "how long is it?" may not make sense.

With a movie, on the other hand, the running time for the film is a key element and moviegoers, unless they walk out in the middle of the film, will experience that actual duration. Movies have directors and producers, screenwriters, and hundreds of other participants from makeup artists to caterers. Some directors are famous, but what makes movies *The Movies!* are the stars: the people that you see on the screen. Having seen the film, most people will be able to relate the story and the names of the primary actors. Very few will remember the producer, although his name will have appeared briefly in big letters at the beginning of the film, and even fewer will have noted the screenwriter's name. The name of the studio that produced the film, analogous to the publisher of a book, is rarely noticed.

All of these above-mentioned creative forms are ones to which its users or participants have a certain emotional attachment. There are other kinds of created resources that we seek out but that are less enticing. I'm thinking of items like census figures, standards documents, technical reports, or court proceedings. If asked about authors of these materials, few people outside of librarianship would name courts or government as authors, although they might see them as responsible bodies of some kind. Users of these materials, however, may be keenly aware of the version of the material; a 1950 census is obviously not the same as a year 2000 census, and a version 0.7 of a standards document would be expected to differ from the 2.0 version. Having the latest version may be essential for some functions, although comparisons of figures across time make use of different versions of the data. Knowing that the copy that you have is authentic and has not been altered is another consideration for these materials. For, like census or economic data, a key factor is whether it is formatted for possible number-crunching.

The point of this brief walk through the various resource types is this: given how different these resources are, and how different our relationship to them is, making any general statement about the structure or data elements needed to describe all resources for all users of a library catalog is going to be difficult, if not impossible. And yet, that is exactly what we do on a routine basis: we create records that treat all resource types the same, and for only one definition of "user." We also ignore or downplay many of the characteristics that are important for users. We often place the names of film actors, when we provide them at all, in a note field that is barely searchable. We also give technical information about data sets and computer files in a note. We give book readers a place of publication and a number of pages but don't give them a clue to the story that the book holds. ("Mentally ill—Fiction" is a subject heading on *Moby Dick*.)

All of this is to point out how varied is our bibliographic universe, and this is without having looked at the differences among users: from novices to experts, children and adults, beach readers and researchers.

Quite clearly, in terms of bibliographic services, one size cannot possibly fit all.

This illustrates the difficulty we have in defining the fundamental nature of the bibliographic "thing," often called a "work." And it also illustrates that the users are an element in that definition. It provides an argument for a flexible treatment that can accommodate a range of user approaches and needs, perhaps a modular structure that can be modified to place emphasis on different information for different materials and different users. Why shouldn't a search on an author return information about the author, including the author's works? Where was the author born, when did she live, what is she known for? In library catalogs, there is no differentiation between Edgar Allen Poe and Barbara Cartland. This isn't neutrality, it's a lack of information. If an item is retrieved on title, there is clearly more that could be said about it than where and when that particular exemplar was published. We present a copy of Charles Darwin's *On the Origin of Species* with a publication date of 2003 without any further explanation, neither of the importance of the work, nor its own true origins. *On the Origin of Species* is meaningful only if you know what scientific thinking was before Darwin's discovery, and that this book is the beginning text for the entire science of evolution.

All of this is possible, but only if we can make some fundamental changes in our approach to bibliographic description. A new approach presupposes a redefining of bibliographic description from a fixed, immovable block of data to a set of interrelated information units that can be viewed from different vantage points.

The challenge for us lies in transforming what we can of our data into interrelated "things" without overindulging that metaphor. There are indeed things of interest to be defined for cultural heritage and creative objects, but our universe of operation lacks the precision of, for example, financial data, where every point of information is precisely known, or the calculation of tensile strength in the engineering task of bridge building. What we describe is not easily subject to quantitative testing, and the difference between success and failure is hard to measure. We are fortunate that errors in library catalogs rarely result in death of the user, but we are hindered by a lack of knowledge of our effect on learning and culture. In spite of the attempts in the 1960s to convince the world that one could add the word *science* to *library* and gain a modicum of status, describing information resources remains an art.

We do have some cold, hard facts in our data storehouse, but we also have some squishy bits—some areas where we simply cannot achieve the level of precision

enjoyed by science and engineering. Part of the reason for our imprecision is the durability of our inventory. Unlike a warehouse of electronic gadgets, we don't discard last year's product when the latest offerings arrive. Some of us even keep the old, the ragged, and the unused materials. Our material lacks uniformity: we have books without authors, articles with citations to prior works that no longer exist, artworks without titles, and boxes of papers that we have not yet had time to open much less cogently catalog. There are works with authors whose real identity is hidden behind the mask of a pseudonym or a coy phrase like "Kind Gentlelady of Upper Norwich" as a way to evade censorship or skirt social norms, and thus to confound library users. We have parts of things that should be whole: scattered issues of a journal, volume two of a three-volume publication, the left side of a triptych.

Sometimes to be precise about what we have, we should be equally precise about what we do not, yet we may not know what we do not have. Some number of works are permanently lost due to war, conflagration, neglect, and low budgets. Creative works arise in a cultural and social context, and only an omniscient cataloger could place all of the items owned by the library in their proper place in the extended history of human thought. Omniscient catalogers are, however, in short supply.

Because we cannot achieve omniscience, we have to take advantage of the technologies available to us. At the same time, we need to retain a healthy skepticism against any promises that technologies, on their own, will solve all of the problems of connecting today's seekers to the wealth of recorded intelligence (and sometimes lack thereof) that may be available through a library.

This book looks at the ways that we define the things of the bibliographic world, and in particular how our bibliographic models reflect our technology and the assumed goals of libraries. There is, of course, a history behind this, as well as a present and a future. The first part of the book begins by looking at the concept of the "work" in library cataloging theory, and how that concept has evolved since the mid-nineteenth century to date. Next it talks about models and technology, two areas that need to be understood before taking a long look at where we are today. It then examines the new bibliographic model called Functional Requirements for Bibliographic Records (FRBR) and the technical and social goals that the FRBR Study Group was tasked to address. The FRBR entities are analyzed in some detail. Finally, FRBR as an entity-relation model is compared to a small set of Semantic Web vocabularies that can be seen as variants of the multi-entity bibliographic model that FRBR introduced.